NOνA Data Acquisition in the Context of Modern Computing

Peter Shanahan

Fermilab

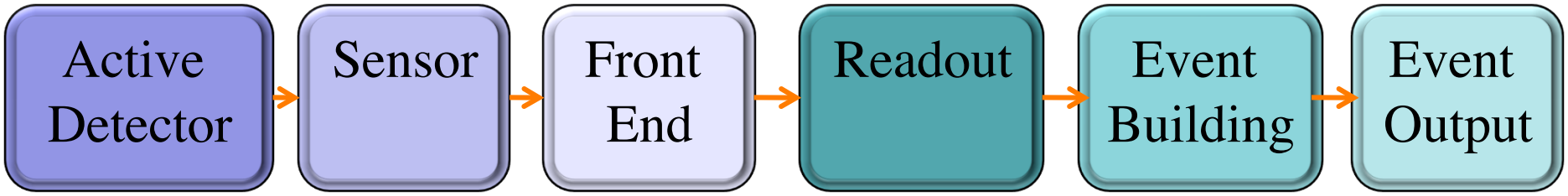
NOνA Summer School

June 20, 2013

Outline

- Data Acquisition Overview
  - With some historical context
- NOvA
  - DAQ Architecture
  - Data flow and formats
  - Control and monitoring
  - Triggering

Typical DAQ System in a Particle Physics Experiment

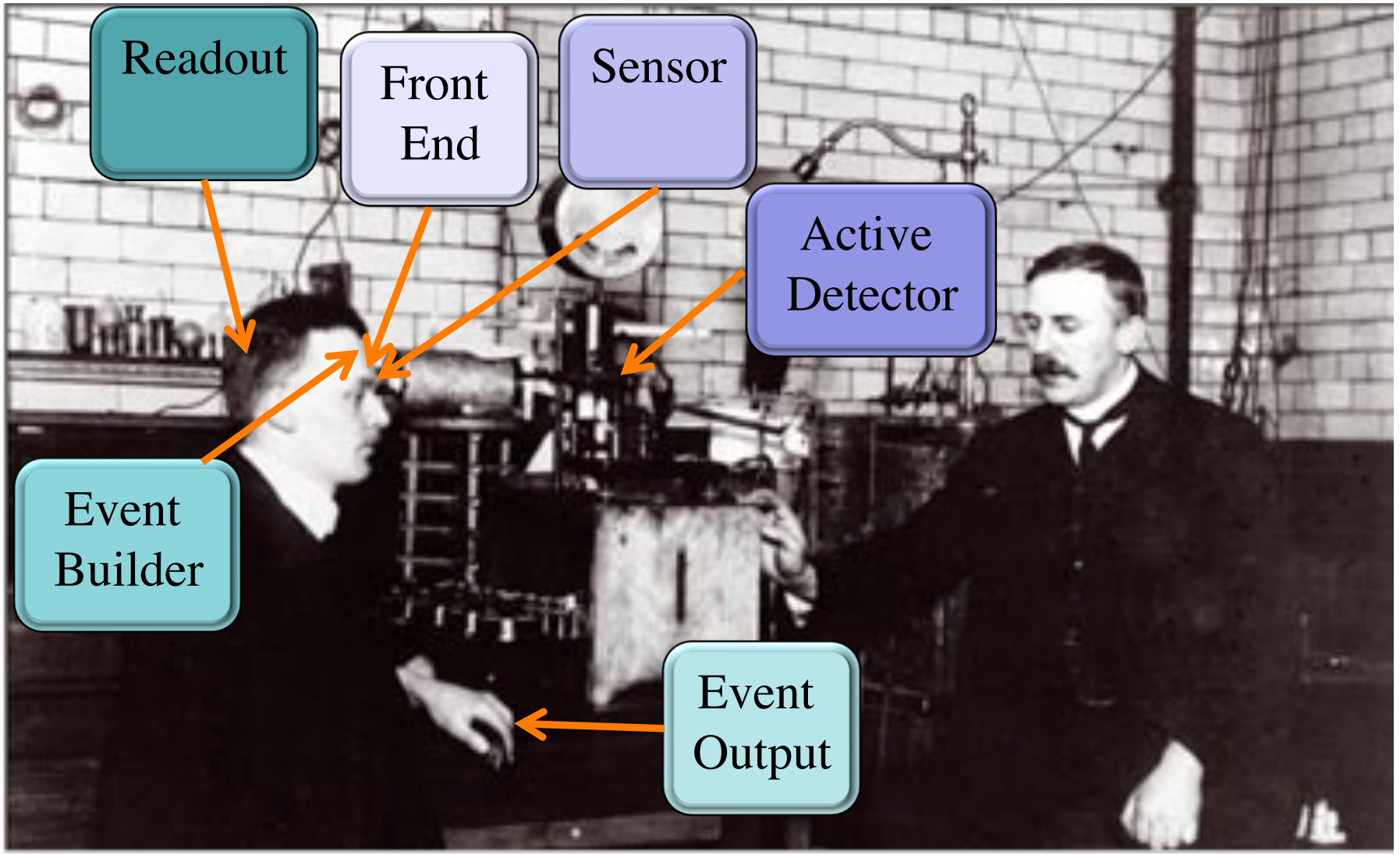


Active Detector	Sensor	Front End	Readout	Event Building	Event Output
Physical signal in response to energy deposition	*Transduces signal from active detector*	*Measures output of Sensor*	*Collects measurements from Front End, transfers to computing space*	*Collates output from readout(s) into Events*	*Saves selected events to persistent storage*

Typical model
Boundaries between stages and other aspects vary among applications

"DAQ" typically refers to these stages

Example

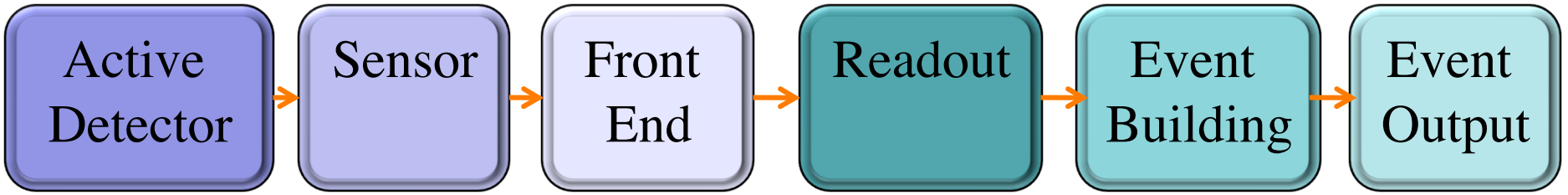

Readout
Front End
Sensor
Active Detector
Event Builder
Event Output

DAQ and Technology

- Particle physics DAQ systems have been a technology driver in the past
  - Fast electronics – *ns* timescale
  - Bus standards - CAMAC, VME, FASTBUS…
  - Computing, including storage
- Still, achieving physics goals with available resources required highly complex systems with clever design
- Example - Triggering

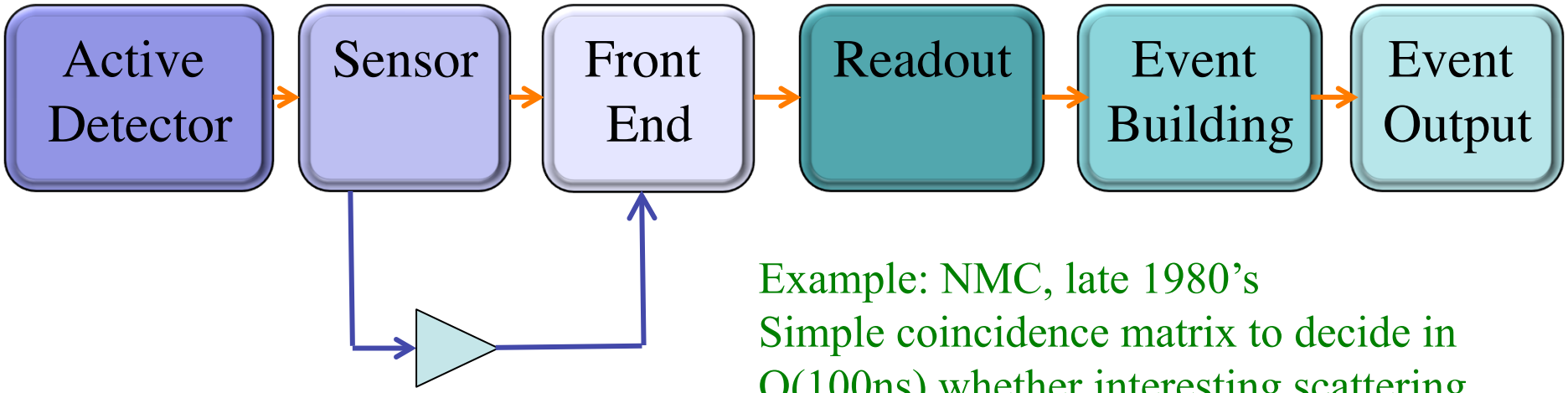
Event Selection

- Triggering
  - Each stage has limited processing and transmission capabilities
    - Processing/reading out events incurs deadtime
  - Triggering: reduce processing of less desirable events to stay within system capabilities



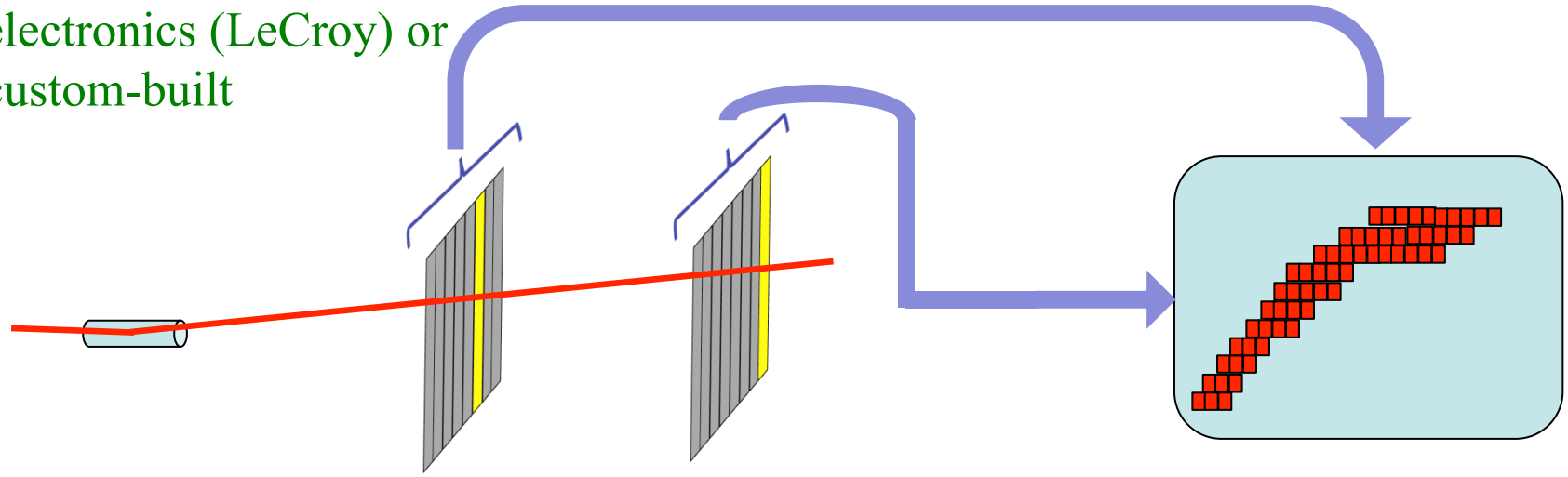
- Example: 3-Level trigger typical in 1980's- 1990's
  - Level 1 – limit Front End processing
  - Level 2 – limit Readout
  - Level 3 – limit Event Output

Level 1 Trigger

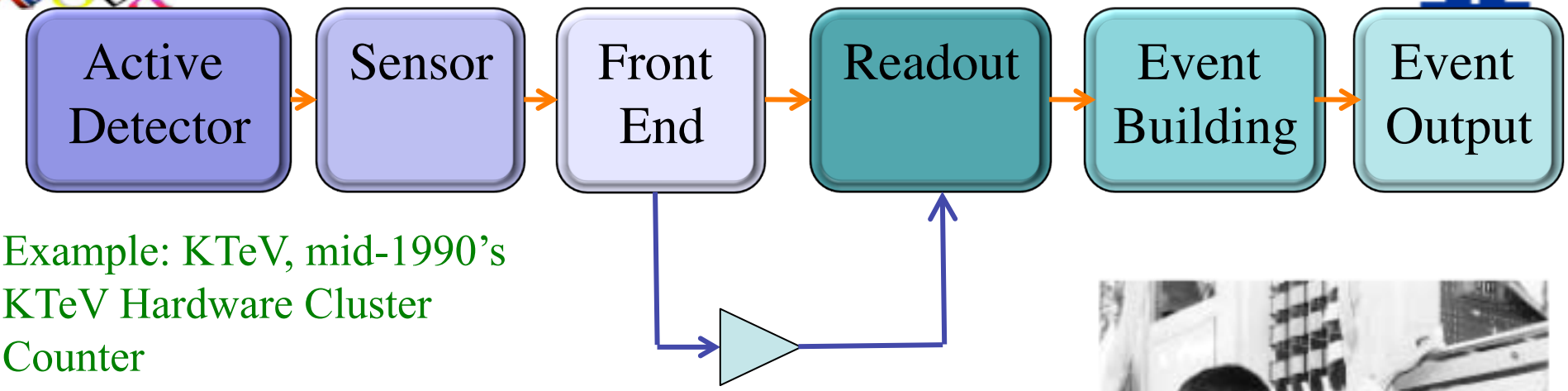


Example: NMC, late 1980's
Simple coincidence matrix to decide in O(100ns) whether interesting scattering angle is present, *before* spending O(10 μs) digitizing LeCroy 2249 ADC

Achievable with standard modular HEP electronics (LeCroy) or custom-built



Level 2 Trigger



Example: KTeV, mid-1990's KTeV Hardware Cluster Counter

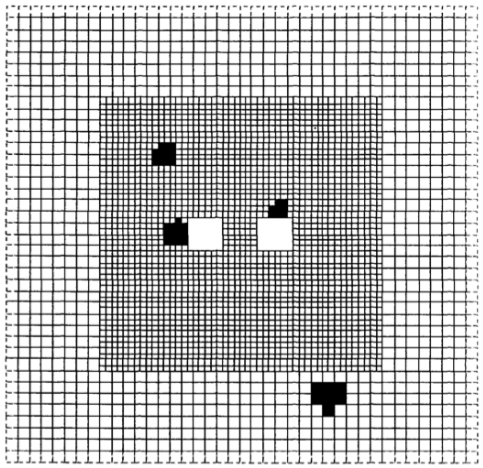
-3600 CsI crystal Calorimeter

-Need to enhance (rare) K_L to $\pi^0\pi^0$ decays while staying within L3 Trigger Input bandwidth.

-Trigger decision in ~1.6μs

Parallelized hardware count of Σ(left turns – right turns) = 4x number of cluster

C.Bown et al., Nucl. Inst. Meth. A 369, 248 (1996)

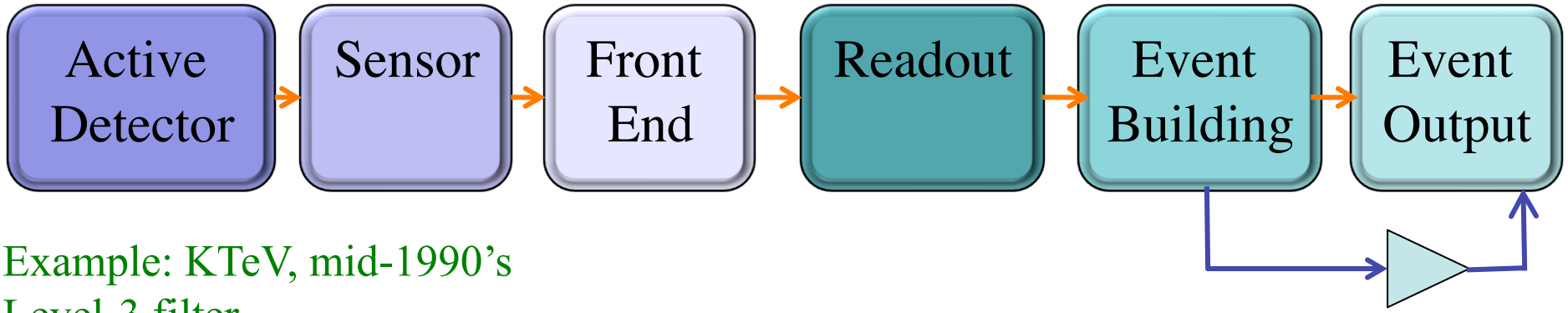


Hit Block Patterns	Pattern Value
	0 - No Turns
	+1 - One Right Turn
	-1 - One Left Turn
	+2 - Two Right Turns
	0 - No Turns

Elliott Cheu works with KTeV hardware at Fermilab.

Good example of a single-purpose custom computer

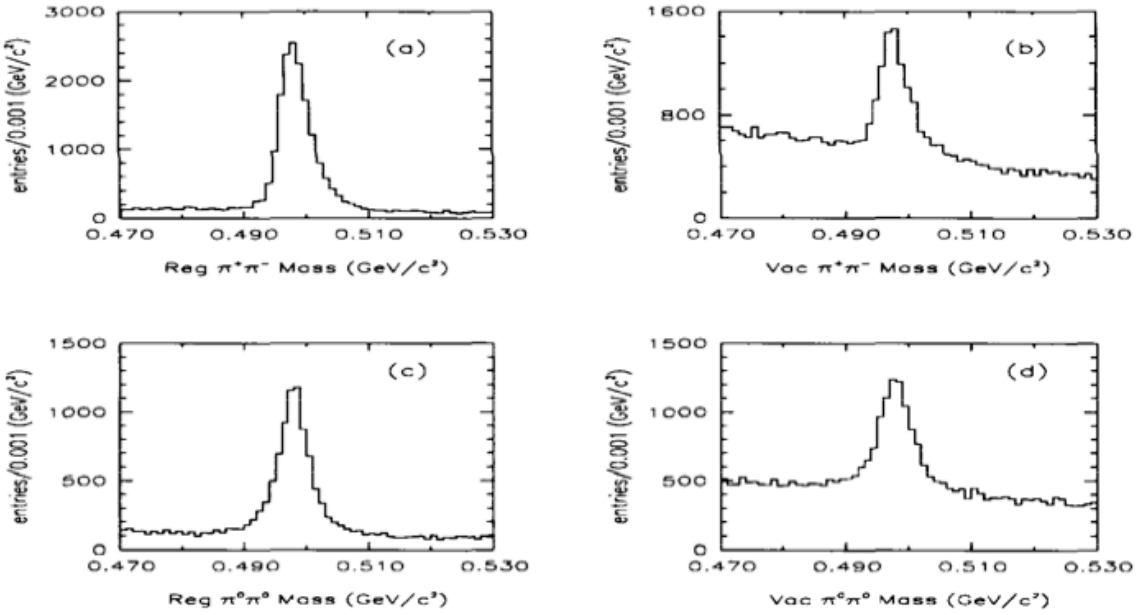
Level 3 Trigger



Example: KTeV, mid-1990's Level-3 filter

One of the first experiments to implement full online reconstruction

Allowed selection of large data set of rare (e.g., K_L - $\pi\pi$) and extremely rare (e.g., K_L - $\pi^+\pi^-e^+e^-$) decays



FERMILAB-CONF-96-384-E, Whitmore, J. (KTeV)

Last Decade – Computing and Networking catches up

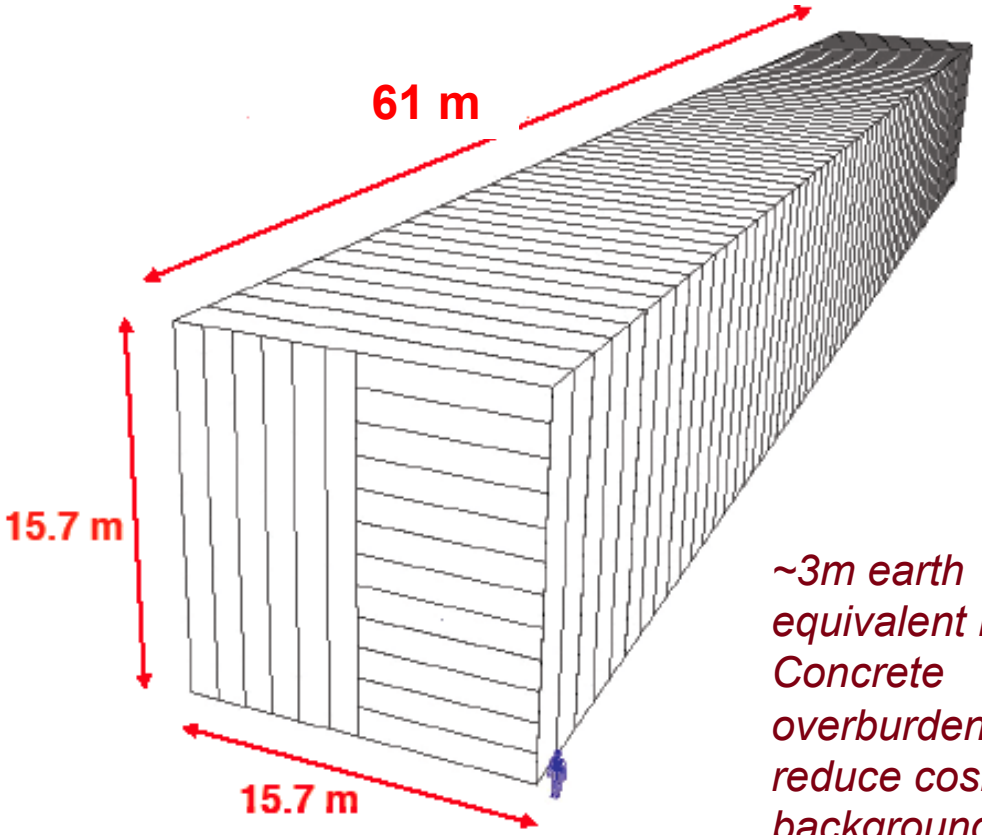
- "Cheap", high rate, gigabit Ethernet
  - Replace expensive and/or fragile and/or proprietary data cables
- Cheap, high-performance, commodity computing
- Cheap high-volume disk space

Year	1985	1995	2010
Machine	VAX	SGI Challenge	Commodity pizza box
Experiment	E665	KTeV	NOνA
Cores	1	8	16
Clock Speed (MHz)	5	200	2660
RAM (MB)	4	2000 (total?)	2000/core

Commodity Computing in DAQ

- For many applications, move away from custom data networks and high-end computing
- Replace with commodity computing and networking
- For example, NOvA…

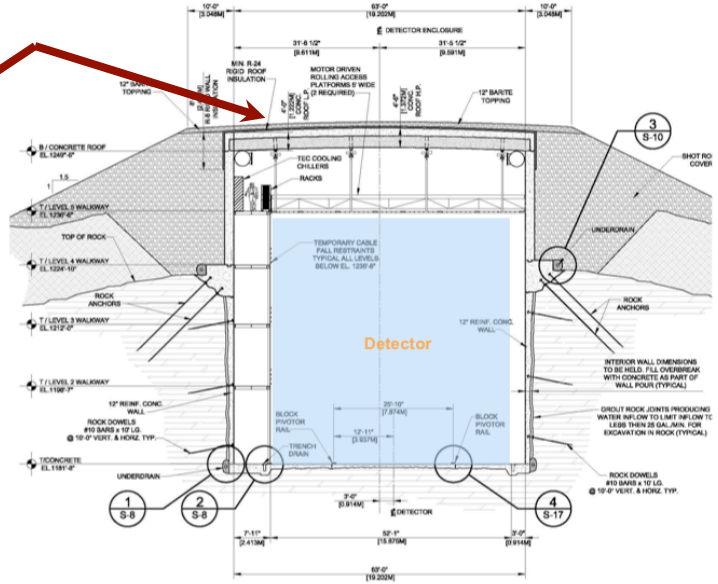
NOvA Far Detector



Ash River, Minnesota

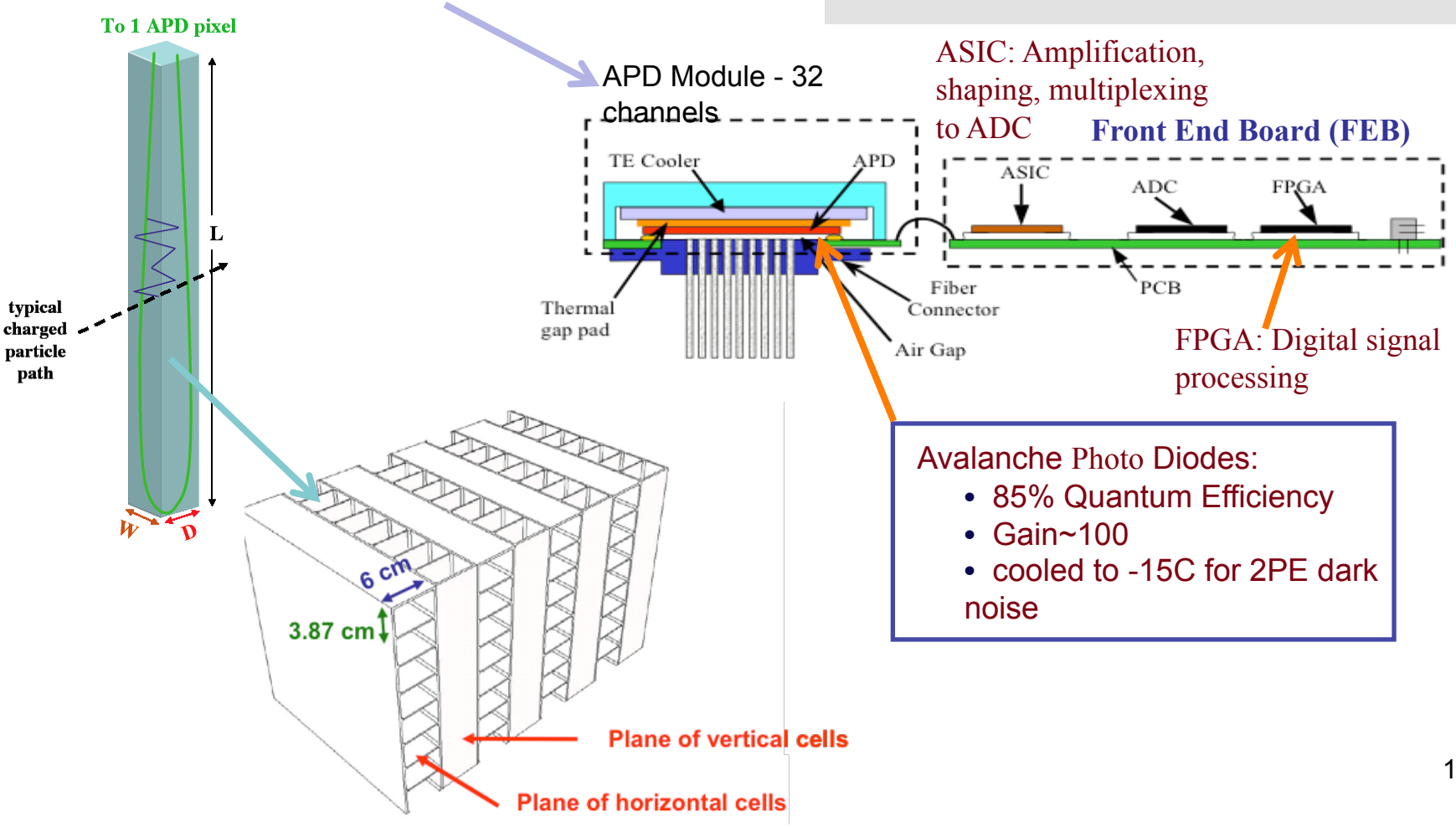
- Liquid Scintillator in PVC structure
-14 kT total mass
-896 planes supported in blocks of 32
-334064 cells
-10752 readout channels

~3m earth equivalent Barite/ Concrete overburden to reduce cosmic backgrounds



NOvA Sensor and Front End

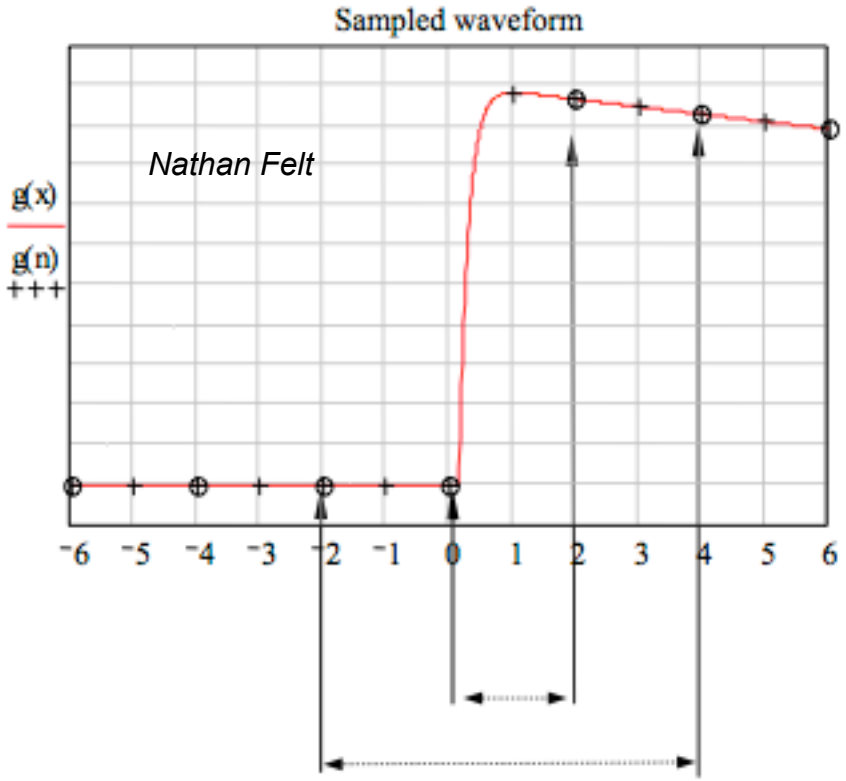
Active Detector → Sensor → Front End → Readout → Event Building → Event Output



Avalanche Photo Diodes:

- 85% Quantum Efficiency
- Gain~100
- cooled to -15C for 2PE dark noise

FEB Signal Processing



ASIC Output Pulse Shaping (FarDet)
-380 ns rise time, 7 μs fall time
-Sample each channel every 500 ns
-Actual clock runs at 62.5 ns, but four 8:1 multiplexors are used to handle the 32 pixels economically

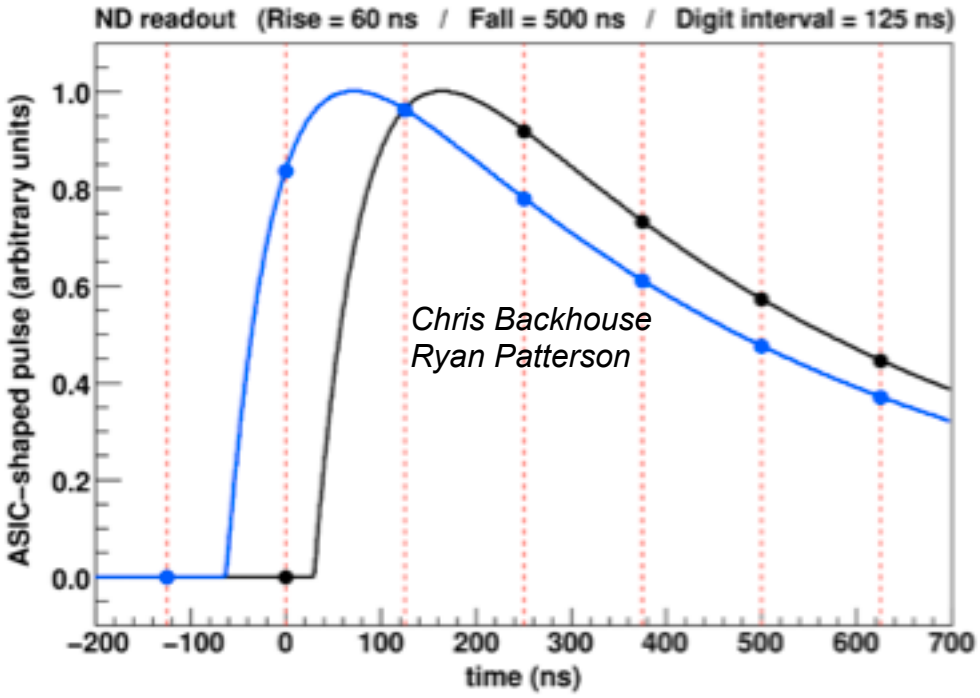
NearDet
-60ns rise time, 500ns fall time
-Sample each channel every 125 ns
-Same base clock, but 2:1 multiplexing

DSO Mode
-Digital Sampling Oscilloscope
-Take (1000) contiguous samples
-Use to measure noise level and set appropriate DCS threshold

DCS Mode
-Dual Correlated Sampling
-Output data when V_i-V_{i-2} exceeds configured threshold
- Data: pulseheight = V_i-V_{i-2} , timestamp=T_i

Fast Timing

Two pulses ~93 ns apart with identical DCS output Pulseheight and Time



Multi-point sampling

-Apply DCS-threshold, but read out multiple contiguous samples

-With compression, can read out up to 12 samples with only 2 addition data words

-Apply matched filtering to improve timing resolution substantially

- -Factor of 3 improvement demonstrated in Data with 4 points

Benefits

-Reduce Near Detector Pile-up

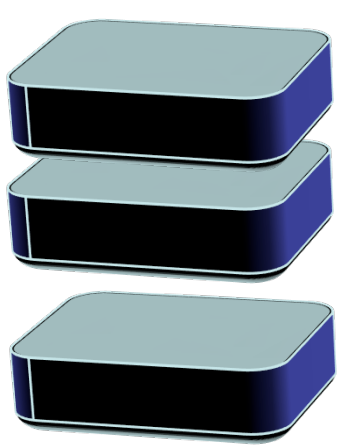
-FarDet: track direction from timing

-etc

DAQ Requirements

- Record beam spills in a window of at least 30 μs, with no dead time
  - Every 1.33 seconds
- Record ~100 times as many periodic cosmic spills with no deadtime
- Cosmic rays (200 kHz!) and noise add up to a hit rate of about 52 MHz, or 620 MB/sec
  - Process 620 MB/sec readout with no deadtime
  - Write out 1.4 MB/sec saved events.

NOvA DAQ Architecture

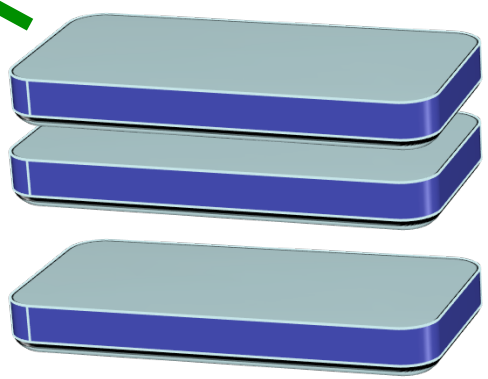


10 gigabit Ethernet switch

Data Logger – Event Output
- Like Buffer Nodes, with
-22 Terabyte of RAID array

Data Concentrator Modules (DCM) – Readout
-PowerPC-based custom computers
- Read out 64 Front End Boards each
- Sends all hits over threshold to Buffer Nodes
- 168 to cover 14 kT detector

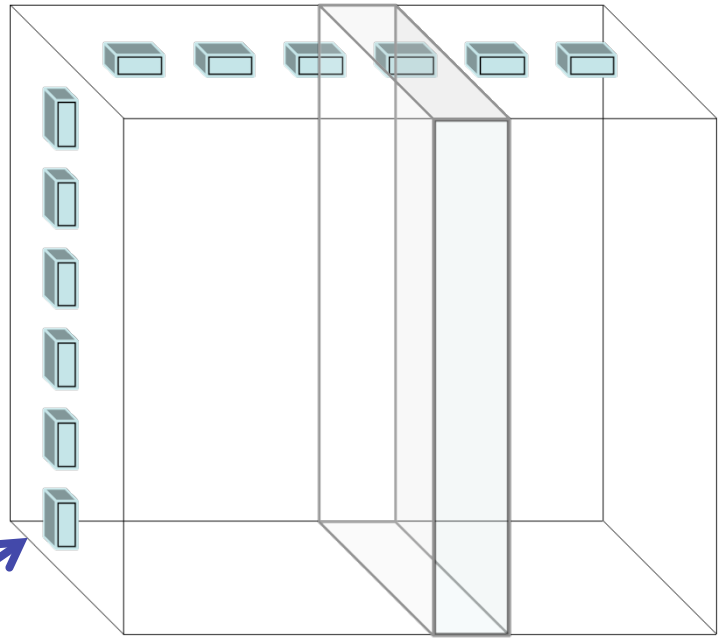
Buffer Farm – Event Buffering and Building
-O(100) Commodity Linux nodes
-16 cores, 2.66 GHz, 2 GB RAM/core
-<$1000 node
-Stores every hit above zero suppression threshold for more than 20 seconds
- Collates data from DCMs
- Sends triggered data to Data Logger



NOvA DAQ Layout

DAQ Cluster (Buffer Farm and Control Nodes)
Ash River – Computing center in Detector building
Fermilab – Lattice Computing Center

168 DCMs for the Far Detetor
620 MB/sec readout
~4 MB/sec/DCM

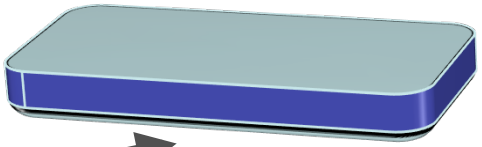


1 Diblock :
12 DCMs mounted on detector
6 per view
Each DCM reads out a 2-module-wide slice

NOvA Trigger Architecture

Trigger Inputs
- Neutrino beam spill gates from ACNET
-Random/Periodic gates
-Soon: Data-driven triggers

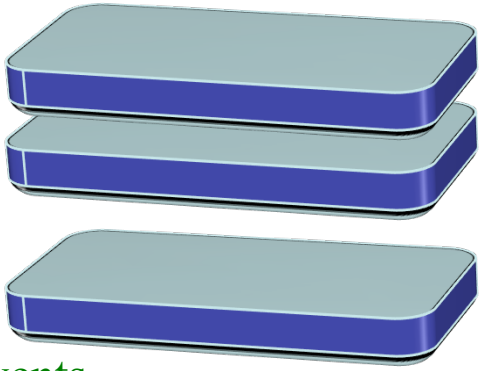
10 gigabit Ethernet switch



Data Logger
- Receives trigger as cross-check on data from Buffer Farm

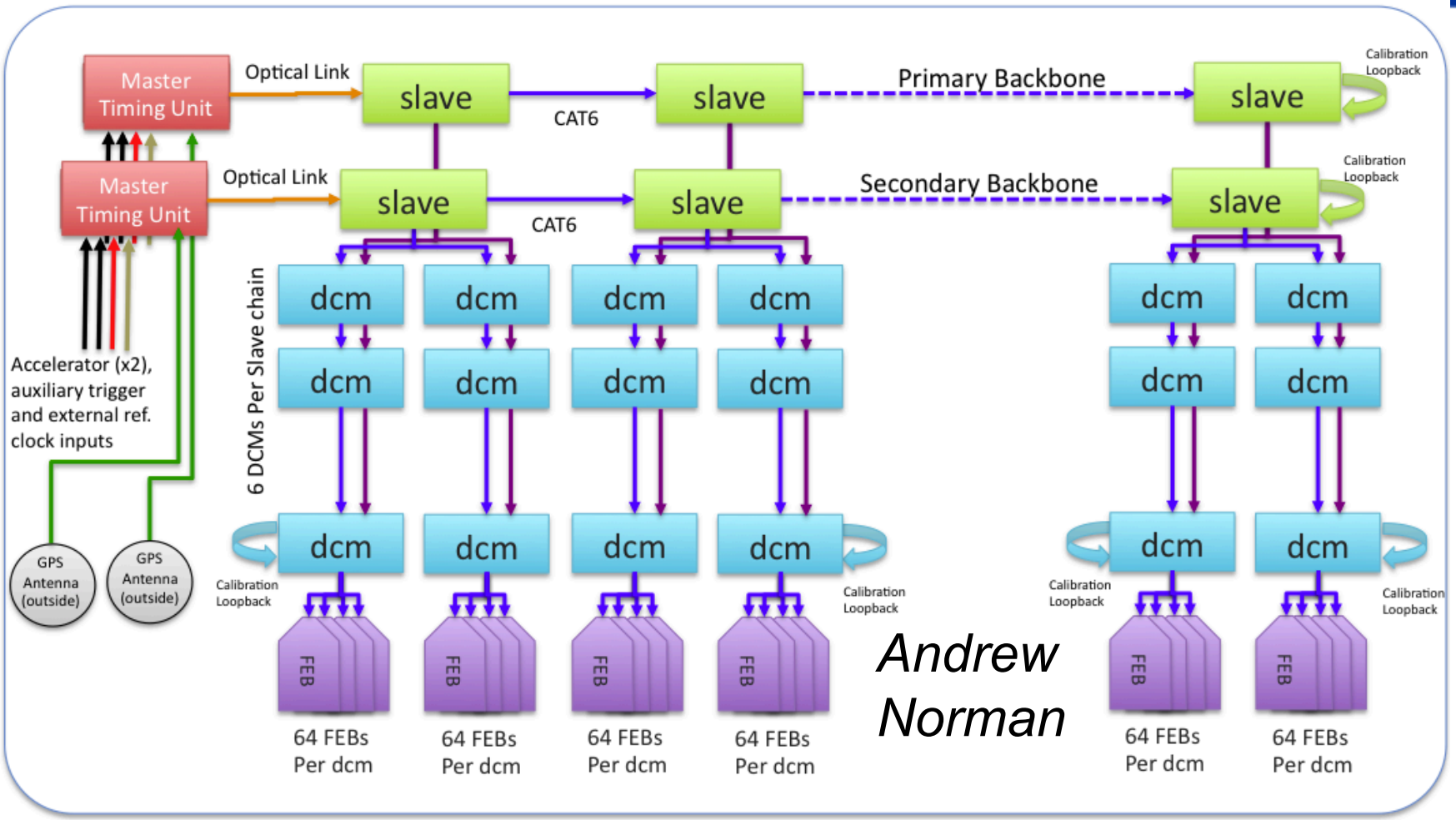
Global Trigger

Buffer Farm
- All Data touching trigger window is sent to Data Logger
-Data-driven trigger:
  -Desirable cosmic events
  -Neutrinos
  -Supernovae
  -Exotic signatures



Trigger Outputs
- (T_0, $T_0+\Delta T$) trigger window
- Sent as software message to Buffer Farm and Data Logger
-No Hardware trigger in NOvA!

NOvA Timing System

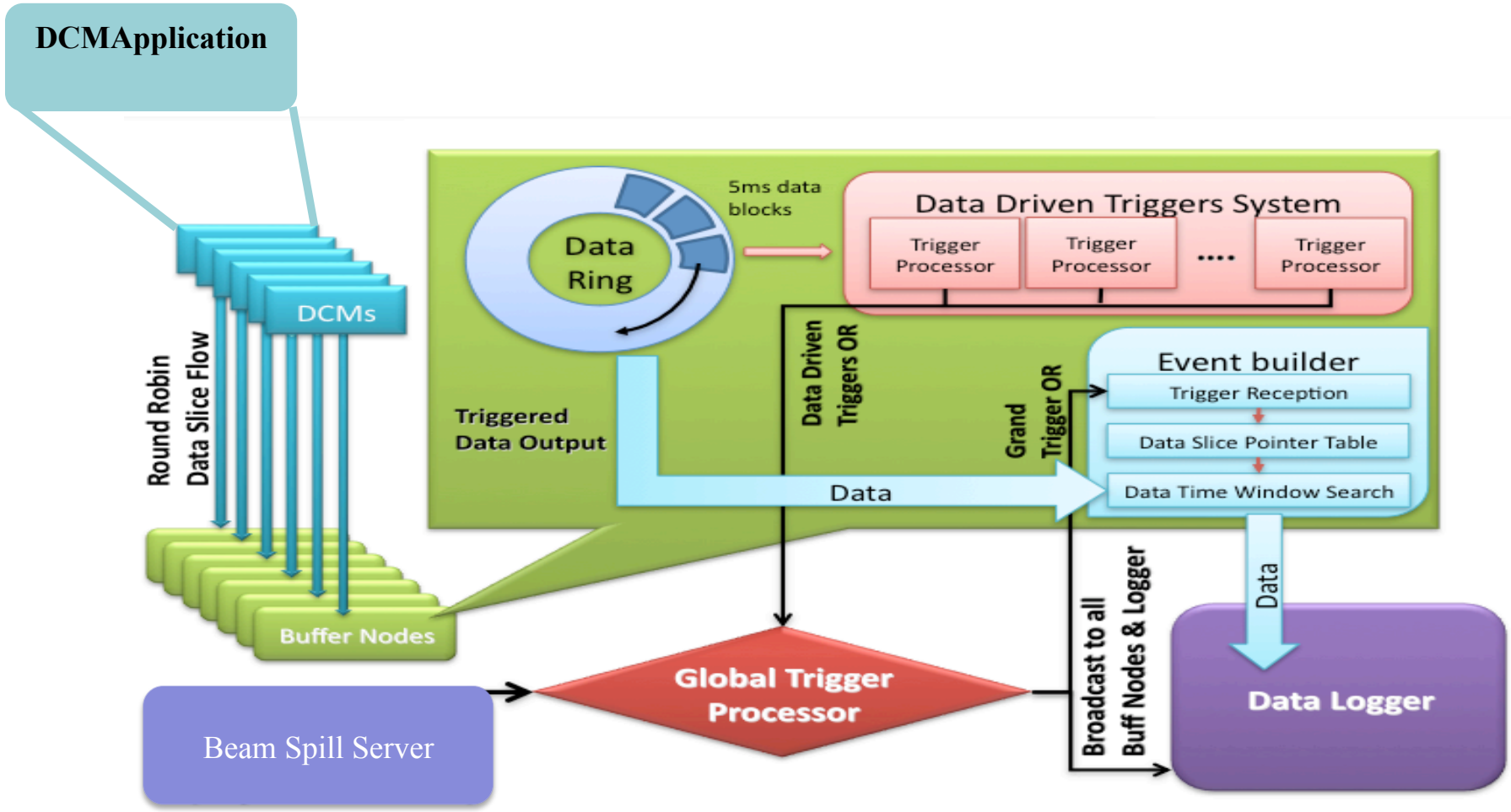


NOvA Clock

-64 MHz clock represented in 56 bits (24+32)

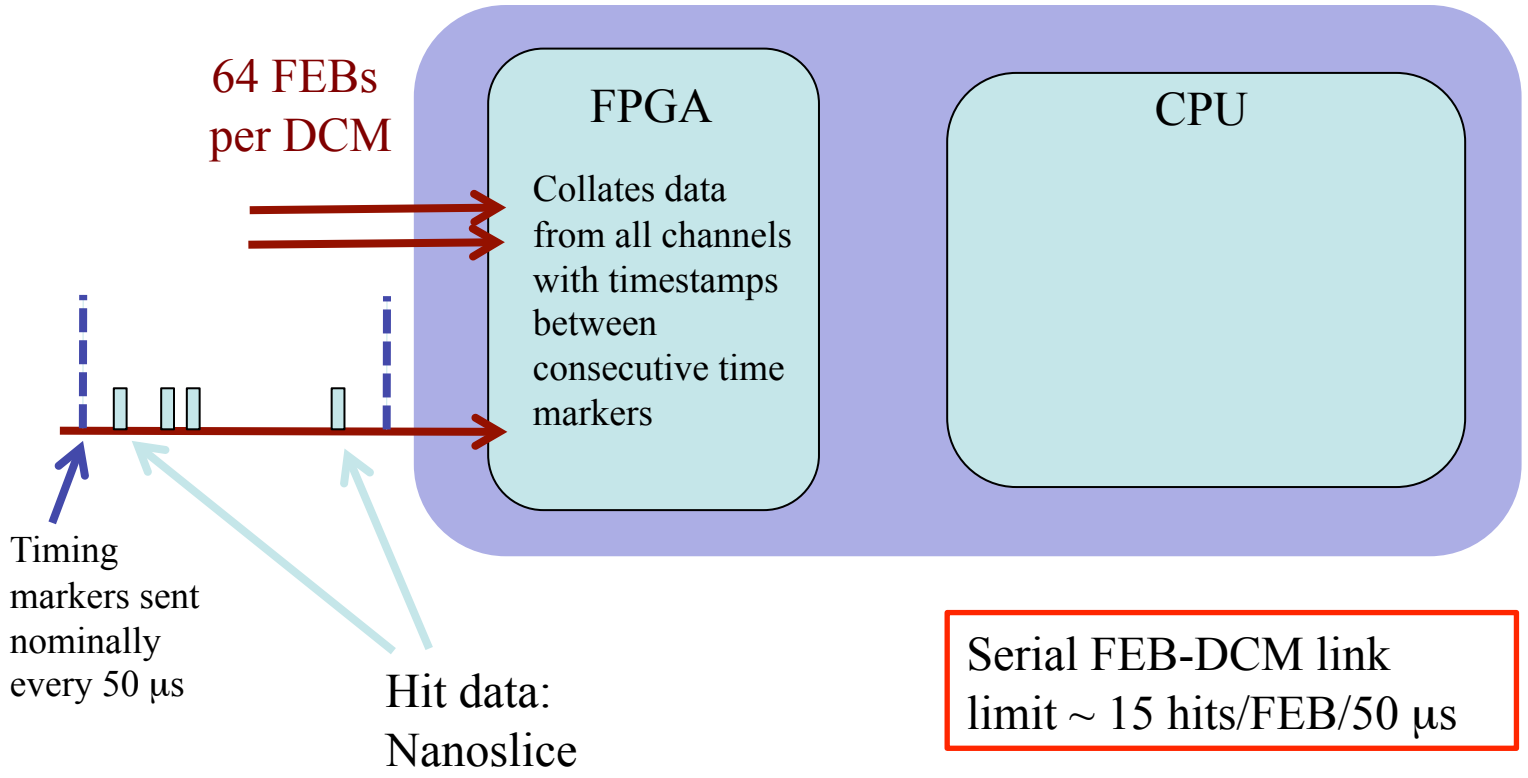
-Covers 36+ year span with 15.625 ns ticks

DAQ Software Data Systems

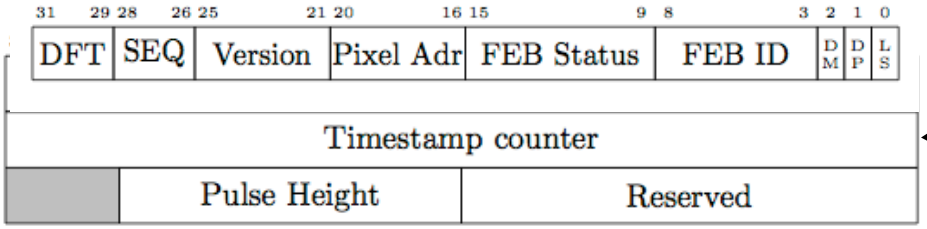


Andrew Norman.

Data Concentrator Module



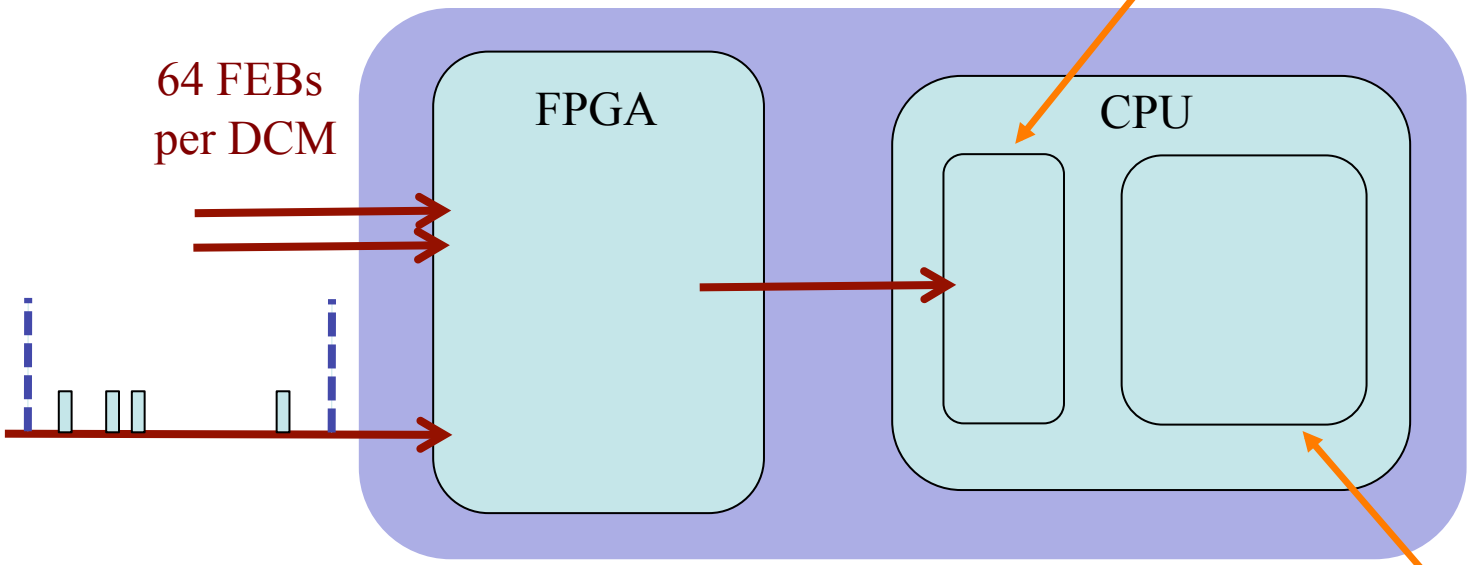
Serial FEB-DCM link
limit ~ 15 hits/FEB/50 μs



32-bit timestamp of 64 MHz clock
Rolls over every ~67 seconds

DCM

"Kernel module"
Serves FPGA data to
CPU as device driver



56-bit 64 MHz time marker
Roll over every 36+ years

Microslice Header
Time Marker Low
Time Marker High
Nanoslice 0 FEB 0
Nanoslice 1 FEB 0
Nanoslice 0 FEB 2
Nanoslice 0 FEB 2
Nanoslice 0 FEB 3

Measured DCM
performance limit:
> 2 hits/FEB/50 μs

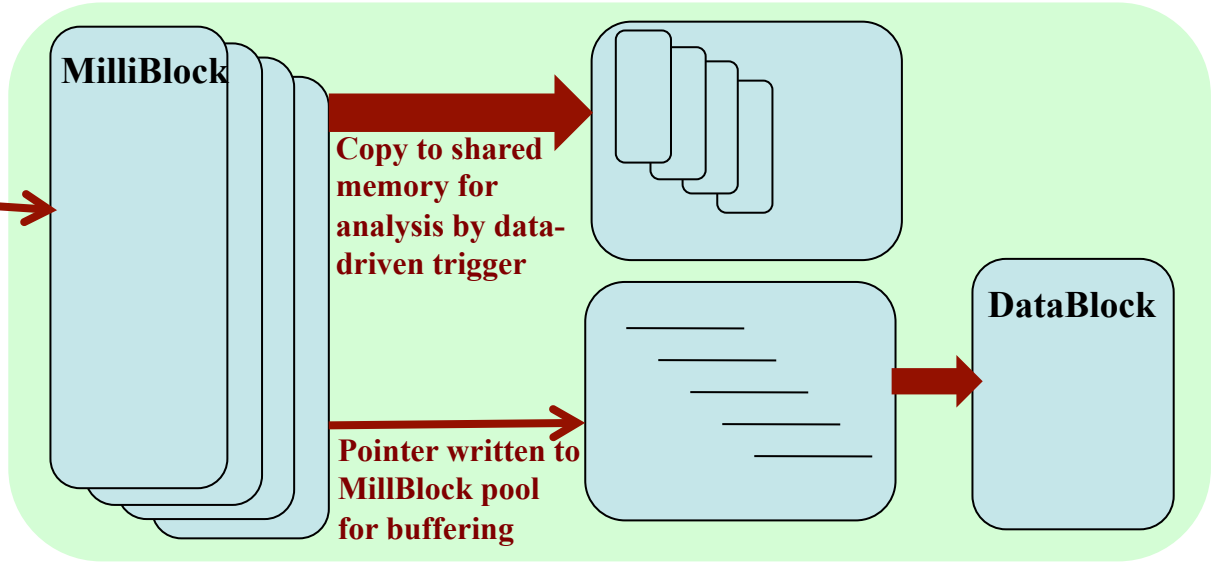
DCM Application
-DAQ system interface for control of DCM and FEB settings
- Monitors data quality
-Produces "Microslices" – data between consecutive time markers
-Builds Millislice for transfer to buffer node: collection of all microslices within 5 ms

Buffer Node

Round-robin input

-All DCMs send to the same Buffer Node for a give 5ms period (sequence number)

- With 100 buffer nodes, this means a buffer node is used twice each second



MilliBlock

-1st millislice seen in a sequence causes creation of new Milliblock

- Complete when last DCM with sequence number is seen

Trigger data search:

- Create a DataBlock for output to DataLogger
- Check all MilliBlocks in buffer
- Does MilliBlock time overlap with trigger time? Check MicroBlocks:
- Copy all MicroBlocks touching Trigger Window into DataBlock

Data Logger

Input

-All Buffer Nodes send a DataBlock in response to a trigger

-For a 500ms trigger window, 90% of triggers will have data on a single Buffer Node

-Write empty DataBlocks, too, for redundancy

Event ("Data Atom" until complete)

-Receipt of Trigger or DataBlock with event number not equal to that of an existing Event will cause new one to be created

- Header, Trigger Info, DataBlocks, Tail

- Complete when data received from Trigger and all Buffer Nodes

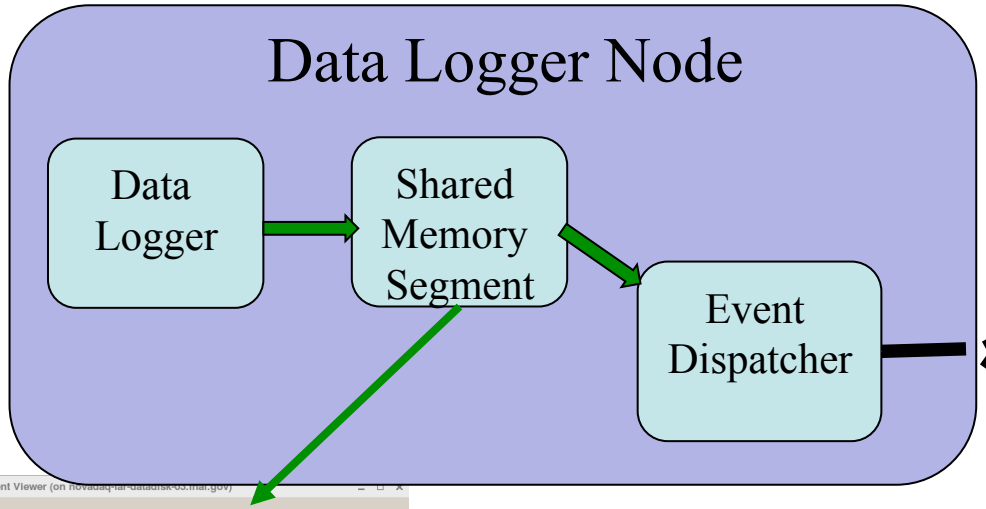
-Note: The trigger, not the hit data, defines the event. The same data can appear in multiple events.

Output

- Send data to configurable set of streams (files on disk) based on trigger ID

-Also send to shared memory segment for monitoring

Data Monitoring



Event Dispatcher servers data to monitoring applications

Memory Viewer

-Displays raw data with highlighting to indicate data block type

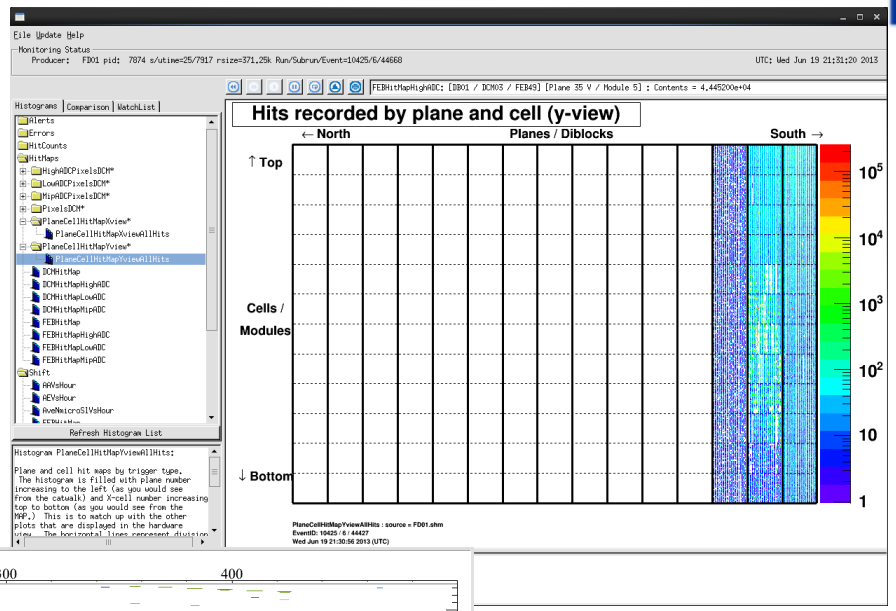
-Very useful for experts

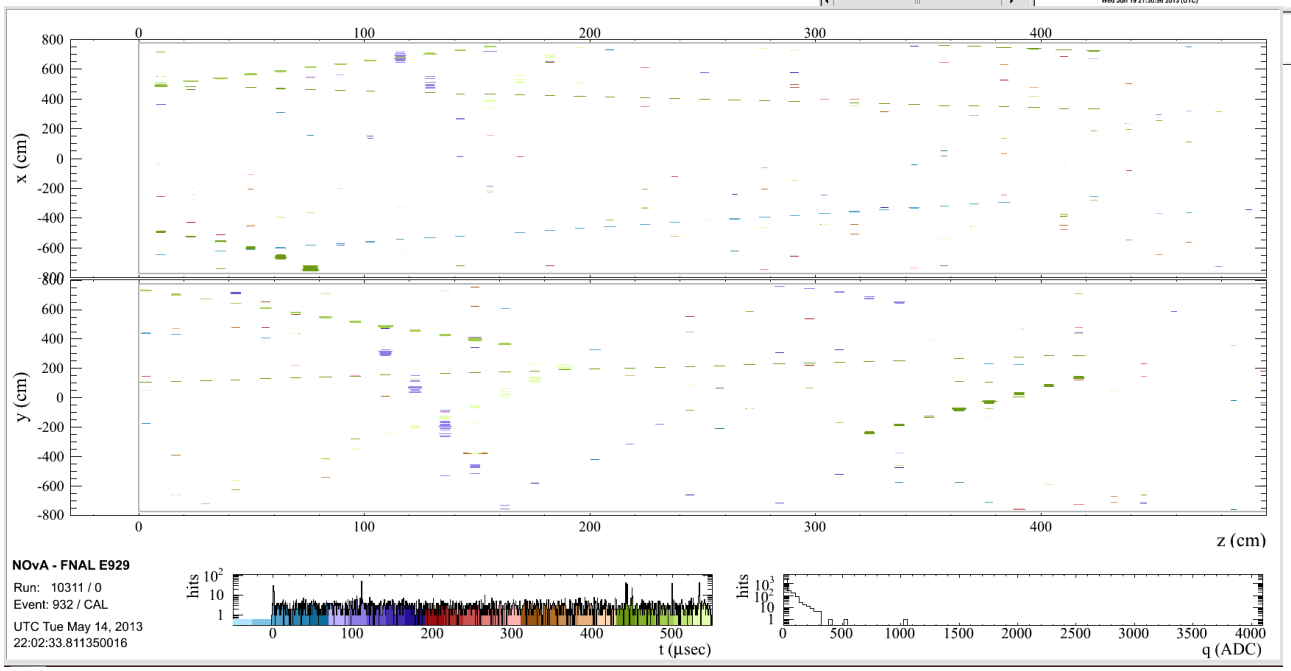
Data Monitoring Applications

ART-based

-Online monitoring: histogram producer/viewer

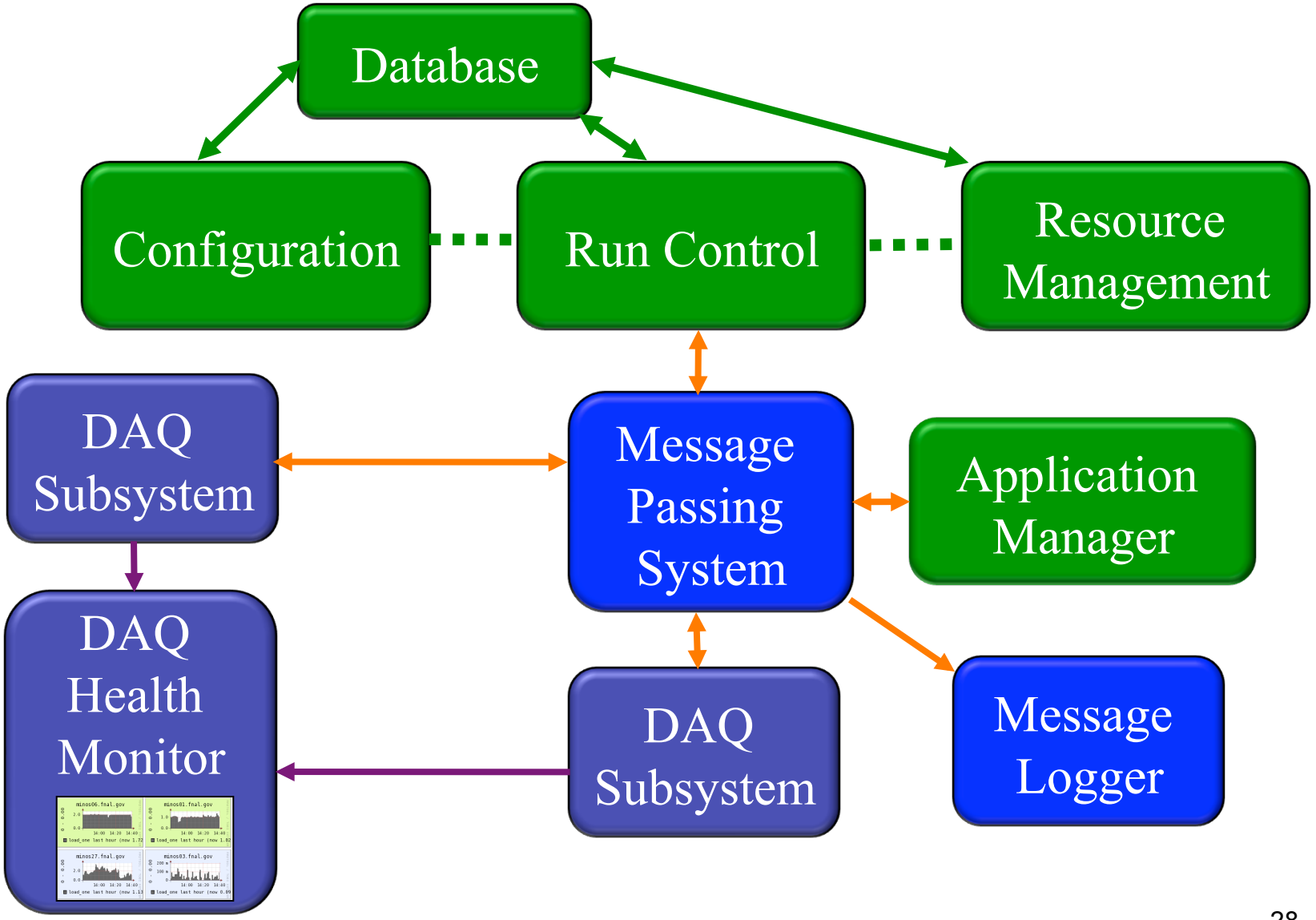
-Event Display: identical to offline version, with different online input stage





Far Detector data

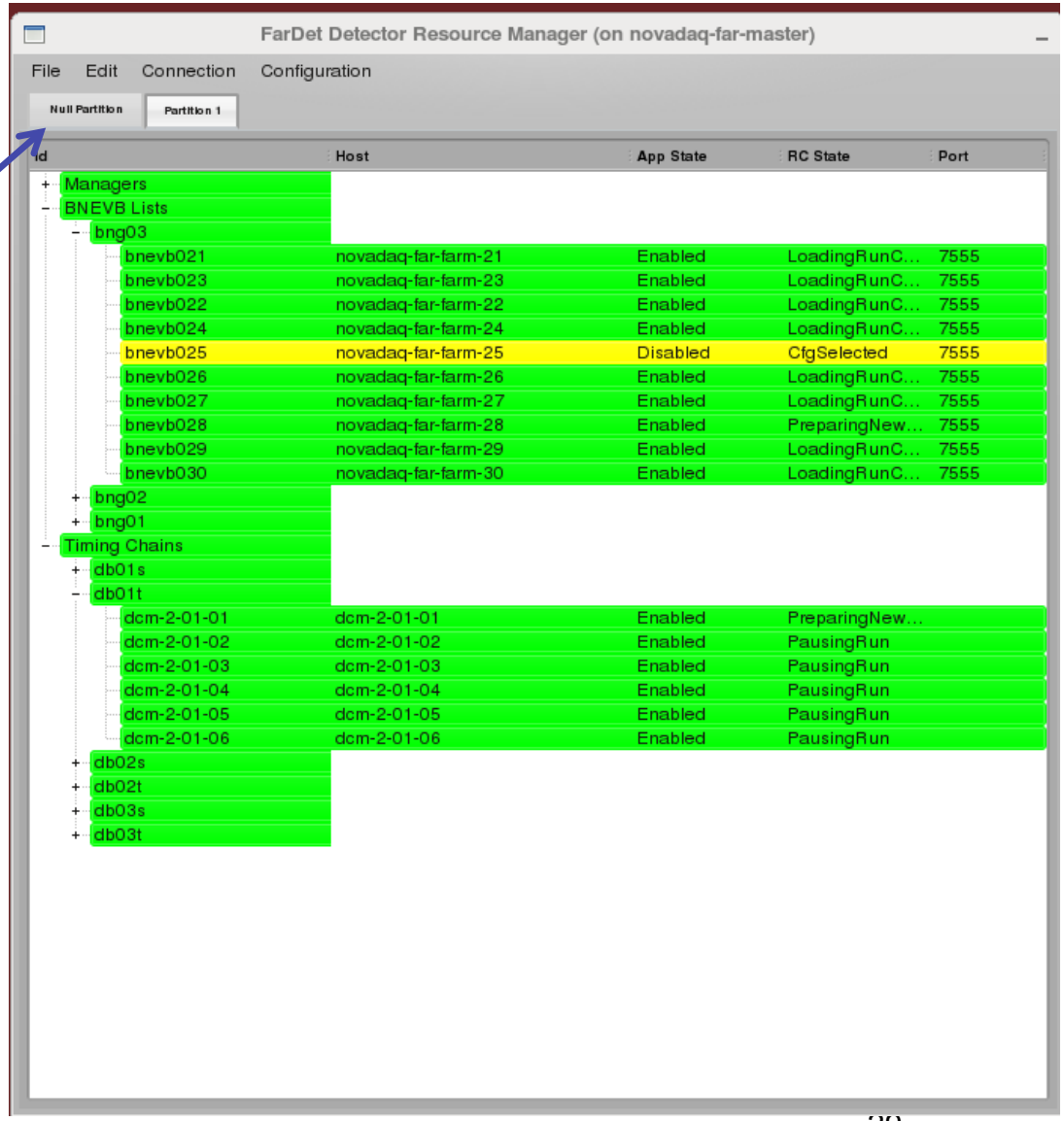
DAQ Control & Monitor Systems



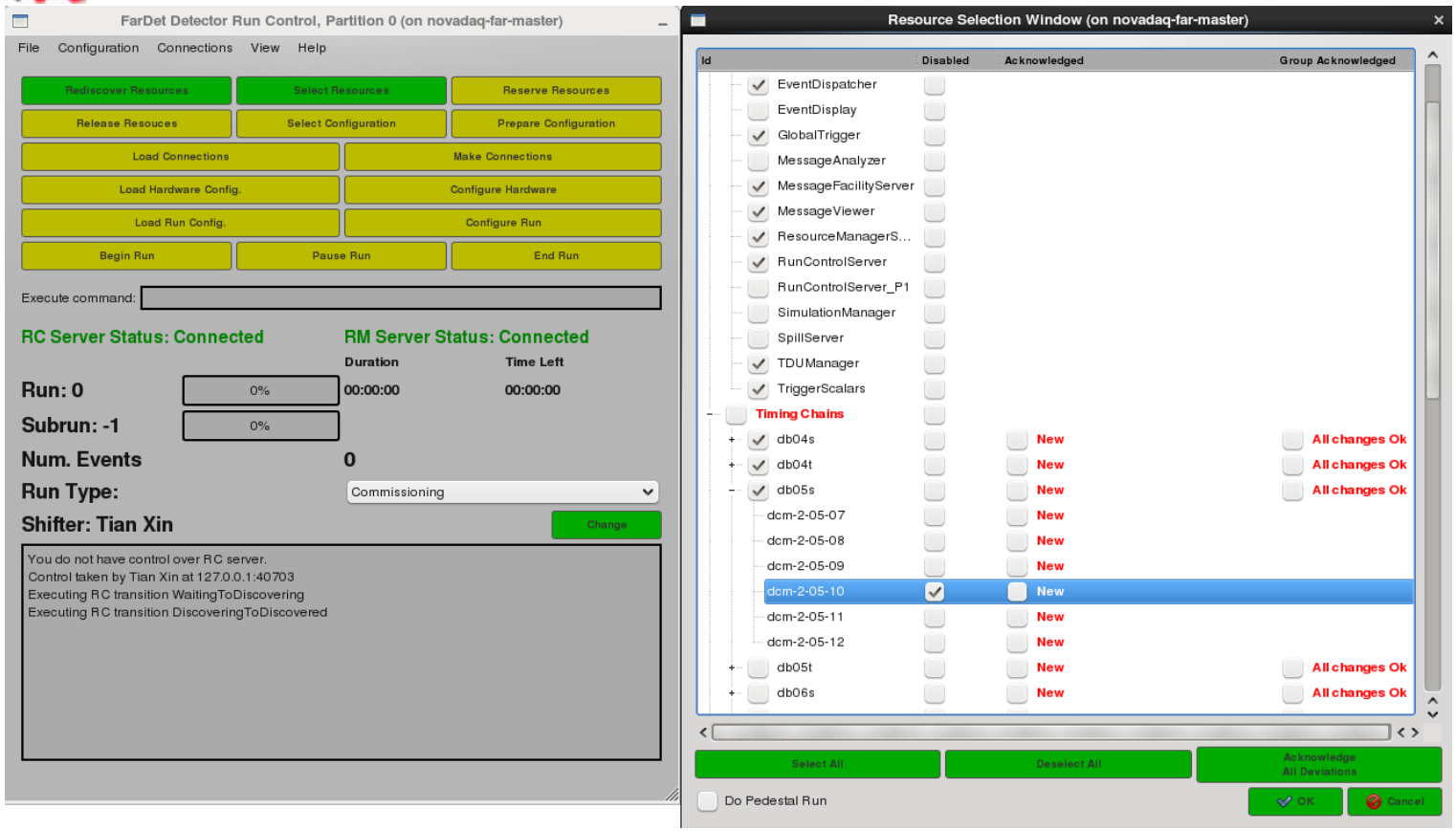
Resource Management

Resource Manager
-Keeps track of which resources are in use
-Resource ~ a DCM or other node

Partition
- A totally independent DAQ system
- One partition can be used for taking physics data on an established part of the detector, while another is used for checkout and commissioning



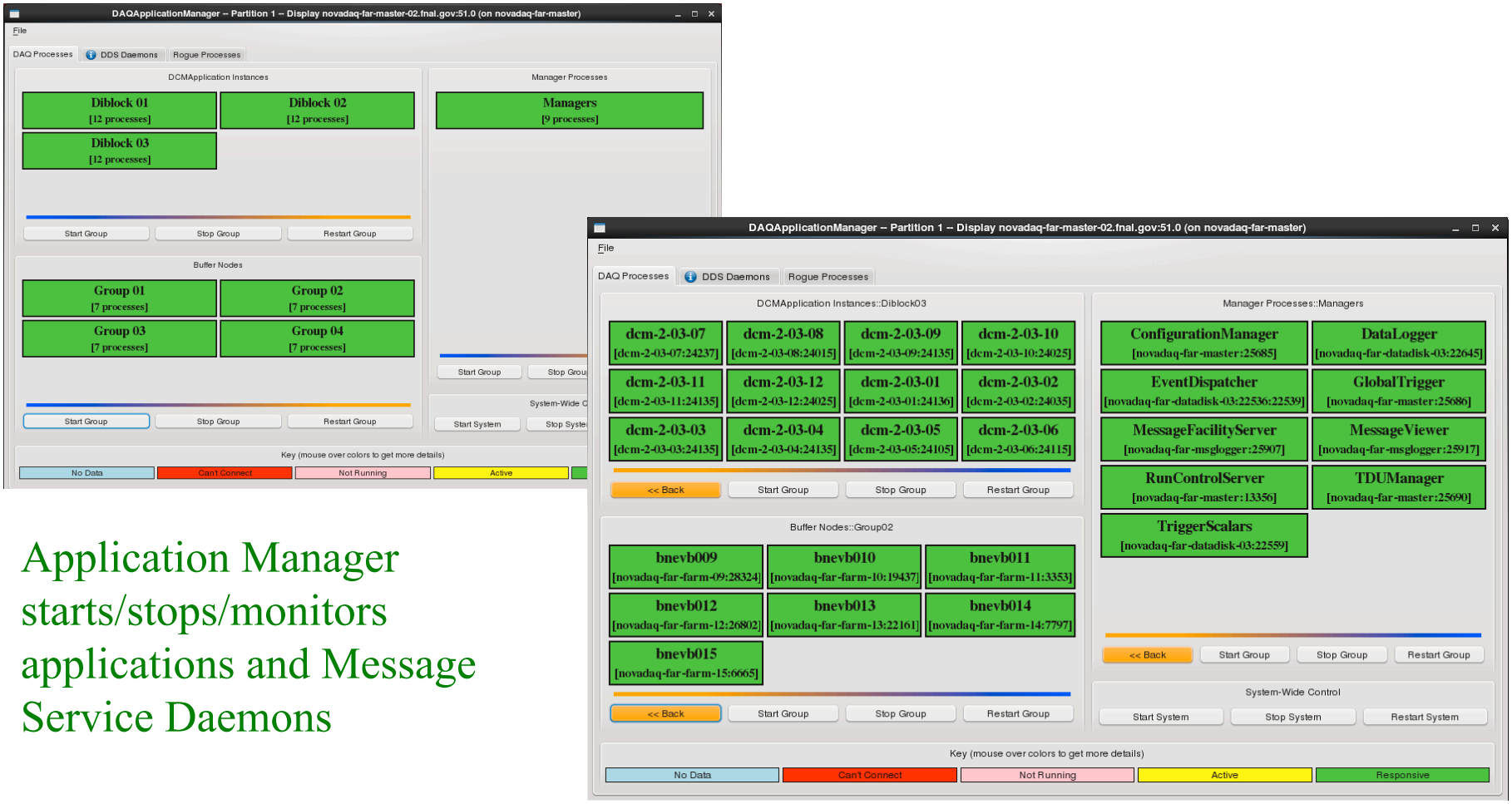
Run Control



User selects resources for partition via Run Control

Once applications are started, Run Control is used to execute configuration steps, and start/stop run

Application Management



Application Manager starts/stops/monitors applications and Message Service Daemons

Message Facility

Message Facility – Log Messages

-Used in offline ART framework

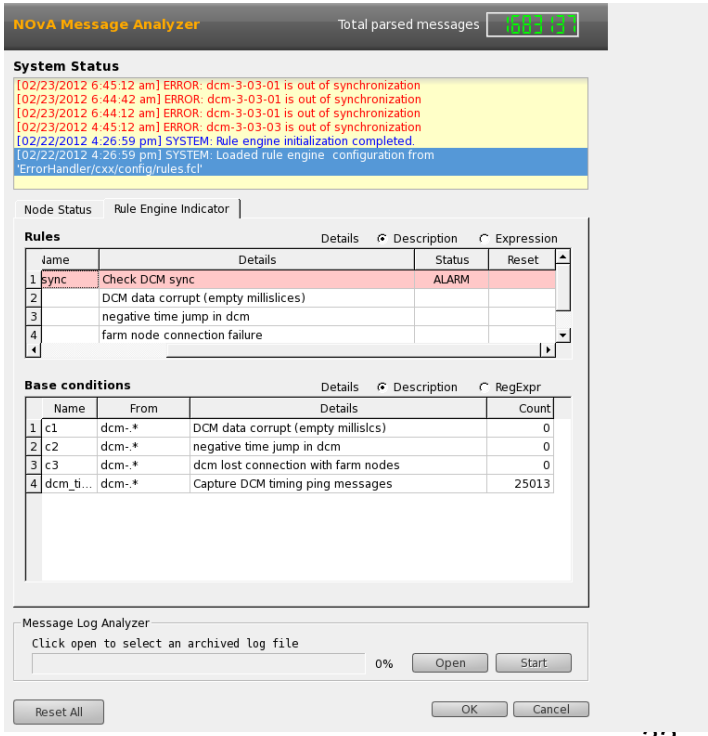
-For Online, has Server destination using DDS backbone

-Message rate throttling is run-time-configurable

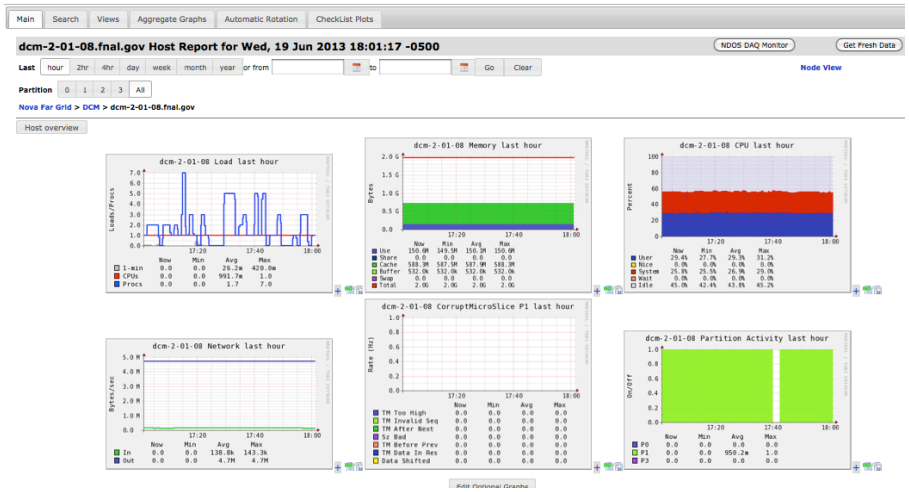
Message Analyzer

-Can be configured to recognize patterns of errors, e.g., many DCMs complain about the same buffer node. Nothing inherently NOvA-specific

-In progress: integration to Automatic Error Recovery actions



NOvA DAQ Monitoring



Ganglia – a widely-used open-source cluster monitoring package

-Daemon on each host sends to central server

-Web-based display of plots of metrics vs. time for "Grid", "Cluster", and nodes

-Default metrics: CPU use, free memory, etc.

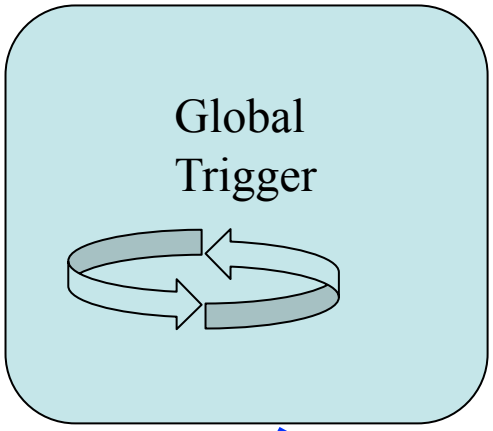
-Extensible to custom metrics

NovaDAQMonitor

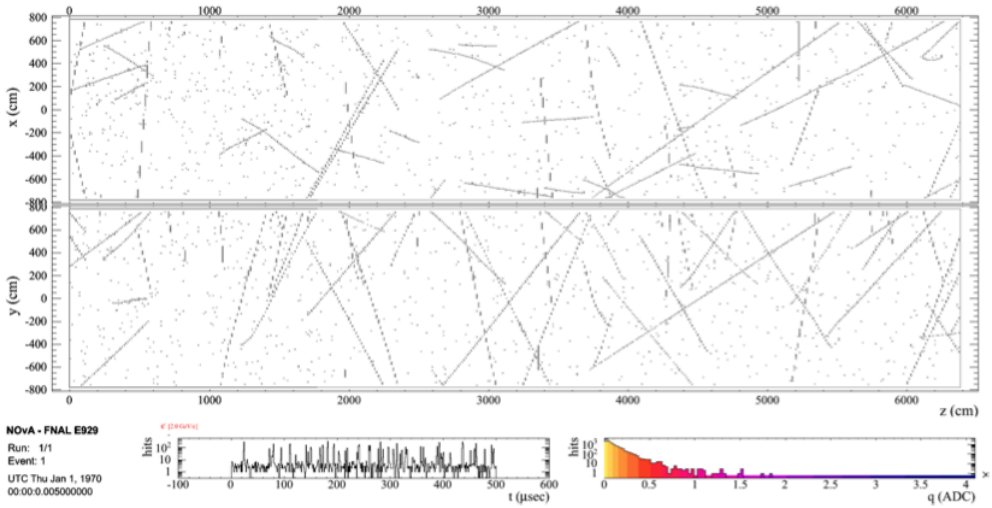
-System for generating warnings and alarms when metrics vary outside configured limits

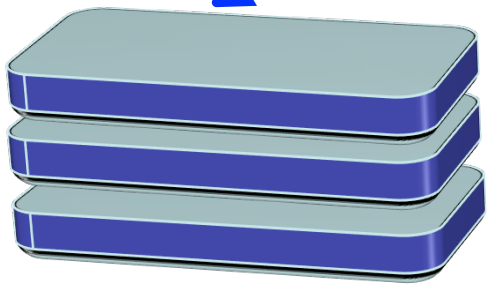
NOvA Periodic Cosmic Trigger

Simple periodic gate
-Configurable rate and width
-Typical:
-10 Hz
-500 μs width



Trigger signals sent to Farm nodes and Data Logger

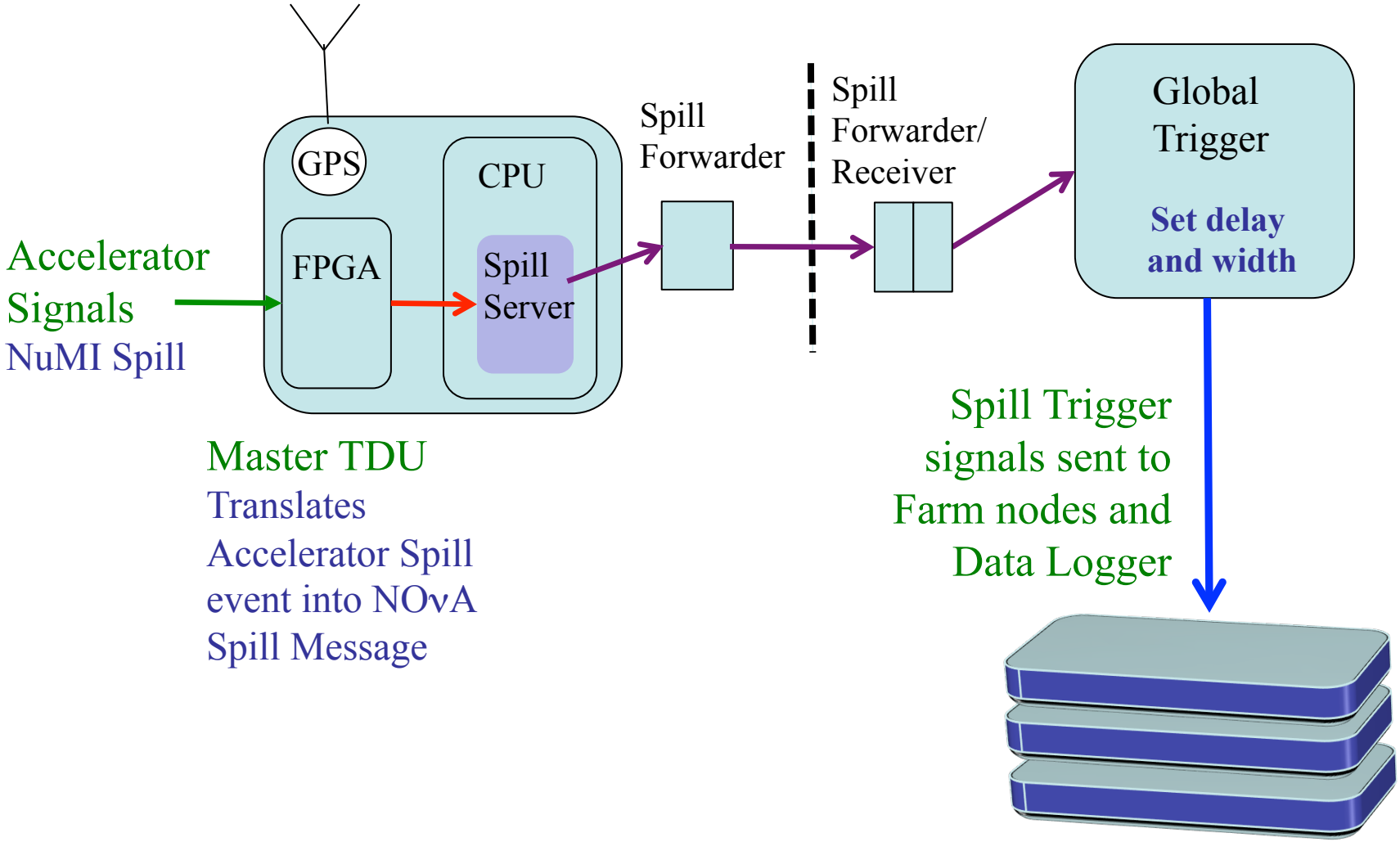




Chris Backhouse

Monte Carlo 500 μs gate

NOvA Spill Trigger



Data-Driven Trigger

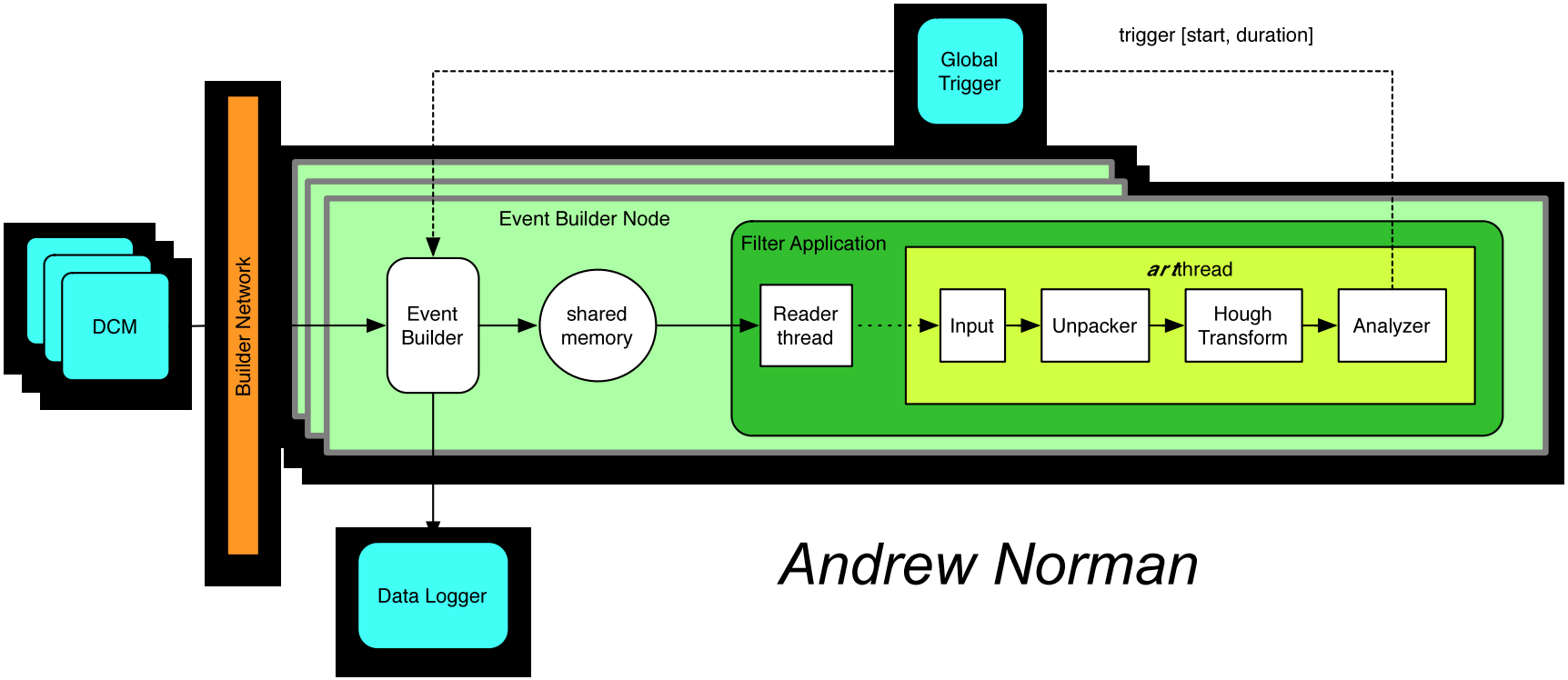
Allows going beyond periodic cosmic gate and spill gates triggers.

-Particularly useful cosmic rays – calibration & alignment

-Supernova neutrinos

-Exotic signatures – e.g., magnetic monopoles

-Testing expected this week



Andrew Norman

Summary

- NOνA Data Acquisition system took Neutrino Data from NuMI and BNB on the NDOS detector
- Far Detector commissioning is in progress
- NOνA is achieving its DAQ design goals with a commodity online buffer farm and commodity network.
- The NOnA DAQ system is the work of too many to acknowledge here
- Special thanks for input and discussions
  - Julie Whitmore, Bob Tschirhart, Vivian Odell, Elliott Cheu,
  - Andrew Norman, Nathan Felt, Chris Backhouse, Ryan Patterson, Ron Rechenmacher, Rick Kwarciany
  - Dan Kaplan
  - And apologies to those I forgot to mention!